Paper C2-1	***USDA Nutrient Data Base for Standard Reference, Release 11***
	David Haytowitz, USDA/ARS/NDL

USDA NUTRIENT DATABASE FOR STANDARD REFERENCE, RELEASE 11

D.B. Haytowitz and S.E. Gebhardt. U.S. Department of Agriculture, Agricultural Research Service, Nutrient Data Laboratory, Riverdale, MD 20737.

ABSTRACT

USDA has been producing electronic forms of Agriculture Handbook No. 8—the USDA Nutrient Database for Standard Reference (SR)—for almost 20 years. Previously the database was only useable on a large mainframe computer; advances in hardware and software have made it possible to do most work with the database on a personal computer. Relational Database Management Systems designed for the personal computer have brought procedures such as specialized queries, data searches, and report generation which once required customized programs, to the desktop and the user. To meet the needs of its users, the Nutrient Data Laboratory has developed a new format for the SR database, using a relational structure. The various files that make up the new SR database and their relationships are described. New data fields have been added to provide more information on the data reported. The layout of these and existing fields will also be described. The structure of the SR database is not the only thing that has changed--many data have been revised and new items have been added. Extensive data on beef and lamb cuts trimmed to 1/8" external fat have been added to the database as have selected new data on ethnic foods. Updated values on breakfast cereals and canned vegetables, are also included in the revised SR database. Data on vitamin E and total dietary fiber have been revised and expanded.

Introduction

The USDA Nutrient Database for Standard Reference (SR) is the major source of food composition data in the United States and provides the foundation for most public and private sector databases. As food composition data are updated, new versions of the database are released. This version, the USDA Nutrient Database for Standard Reference, Release 11 (SR11) contains data on approximately 5600 food items for over 70 nutrients. It replaces the previous release (SR10) issued in July 1993 and adds food composition data for several hundred new items. A number of other items have been updated and new information on all foods such as type of data and INFOODS tagnames (2) have been added. Although most foods, with the exception of breakfast cereals and infant formulas, have generic descriptions, there has been an expansion of data for brand name products. A relational structure was adopted for this release. Where once there was only a nutrient file and a coding manual, there are now several principal files and a number of support files linked by key fields in each file. SR11 includes all the food composition data published in the 21 sections and four supplements of Agriculture Handbook No. 8 (AH-8) (3-27) and supersedes it. In the future, it will be superseded by subsequent releases.

Changes and Updates

Along with changes in structure, there are also changes in the data. New items have been added and old items have been updated. Among these are vitamin E values, which are now all reported in milligrams of a-tocopherol equivalents. New values have been provided for a number of breakfast cereals and those products which are no longer marketed have been removed. Updated data on brand name candies are also available. In the past few years the canned vegetable packers and soup manufacturer's have reformulated their products to lower the sodium

content. The values for sodium in these products have been updated to reflect this change. When the section on baby foods
was first published, no data on infant formulas were included. These also have been added to SR11. New data on beef T-bone and porterhouse steaks have been added. Market basket studies show that the average fat-trim on beef and lamb cuts is now 1/8". In consideration of this change in the market, new beef and lamb cuts trimmed to 1/8" external fat have been added to SR11. In addition, selected items from a contract sponsored by the Nutrient Data Laboratory (NDL) on ethnic foods have also been included. More will be added in the interim release.

SR11 has been made available on the Nutrient Data Laboratory Home Page and Bulletin Board in a variety of formats. Two relational files are available in both ASCII and DBF. An abbreviated file, containing fewer nutrients, is also available in both ASCII and DBF. The data have also been converted to the International Food Distributors Association's (IFDA) Product Data Exchange Format, Version 3.0. (1) which was developed to facilitate the exchange of product information, including nutrient data, between food manufacturers, suppliers and their various clients throughout the food chain. The relational file formats allow users to import and query the data, using a database management software package.

Planning is underway to produce a CD-ROM containing the above mentioned SR11 files. Because of time constraints, the CD-ROM will be available with the first interim release. This interim release will add items, for which data were received too late, to be included with the intial release of SR11. It is anticipated that interim releases will follow about 6 months after the primary release.

An online program is also available on our home page which permits users to look up the nutrient content of any food in the database. The user enters a food name or portion of the name, and is then given a list of items which contain the entered term. After selecting the food item to be displayed the user is given a list of household weights available for that food. The user can select to report the data on the 100-gram basis and up to three household weights, or if the 100-gram option is not selected, up to 5 household weights. The user then gets a report, suitable for printing, of the nutrients in that food calculated to the household weights selected.

The Nutrient Data Laboratories home page has been moved to a web site managed by the National Agricultural Library. The new URL is:

http://www.nal.usda.gov/fnic/foodcomp

Quality Control

A number of tests were conducted to insure data integrity: 1) review of all foods in the nutrient file to determine if there were corresponding records in the description file? 2) summing the proximates to verify that they add up to 100? 3) checking the energy value to see that it is less than the sum of protein and carbohydrate value times four and the fat value times nine? 4) summing the mineral values to verify that it is less than the ash value? 5) checking to see if the vitamin A value reported in RE is less than that reported in IU? 6) summing the total saturated, total monounsaturated and total polyunsaturated fatty acids to verify that it is less than the total fat value? 7) summing of the individual fatty acids to verify that it less than the total fat value? 8) summing the individual saturated fatty acids to verify that it is less than the value for total saturated fatty acids? 9) summing the individual monounsaturated fatty acids to verify that it is less than the value for total monounsaturated fatty acids? and 10) summing the individual polyunsaturated fatty acids to verify that it is less than the value for total polyunsaturated fatty acids? Reports listing any exceptions were reviewed by food specialists and either explained or corrected.

Explanation of File Formats

The database is comprised of several separate files. There are three principal files: Food Description File (Table 1), Nutrient Data File (Table 2), and Gram Weight File (Table 5). There are four support files: Nutrient Definition File (Table 4), Measure Description File (Table 6), Source Codes File (Table 7) and Food Group Description File (Table 9). A diagram showing the relationship between these files is given in Figure 1.

Figure 1 - Relationships between Principal Files and Support Files

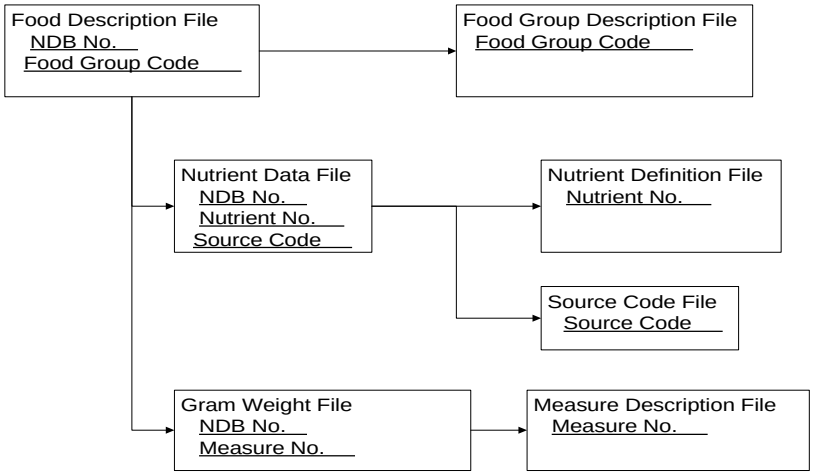

Table 1 - Food Description File: (File Name = FOOD_DES)

The Food Description File contains both a long and short description for 5,635 food items along with the scientific name, refuse, and the factors used for calculating protein, calories and fatty acids if applicable.

- Links to the Food Group Description File by the FdGp_Cd field.
- Links to the Nutrient Data File by the NDB_No field
- Links to the Gram Weight File by the NDB_No field

Table 1 - Food Description File: (File Name = FOOD_DES)

Field Name	Type	Blank	Description
NDB_No	A 5*	N	5-digit Nutrient Data Bank number which uniquely identifies a food item
FdGp_Cd	A 4	N	4-digit code indicating food group to which a food item belongs.
Desc	A 200	N	200 character description of food item
Shrt_Desc	A 60	N	60 character abbreviated description of food item. Generated from the 200 character description using abbreviations in the Appendix. If short description was longer than 60 characters, the food specialist made additional abbreviations.
Ref_desc	A 45	Y	Description of inedible parts of a food item, such as, seeds or bone.
Refuse	N 2.0	Y	The percent refuse
SciName	A 60	Y	The scientific name of the food item. Given the first time the food appears in the file, if applicable.
N_Factor	N 4.2	Y	Factor for converting nitrogen to protein
Pro_Factor	N 4.2	Y	Factor for calculating calories from protein.
Fat_Factor	N 4.2	Y	Factor for calculating calories from fat.
CHO_Factor	N 4.2	Y	Factor for calculating calories from carbohydrate.
Blank_1			Field reserved for future use
Blank_2			Field reserved for future use
Blank_3			Field reserved for future use
Blank_4			Field reserved for future use

Table 2 - Nutrient Data File: (File Name = NUT_DATA)

The Nutrient Data File contains the nutrient values and information about them including, sample count and standard error for analytical values and a source code indicating the type of data.

- Links to the Food Description File by NDB_No
- Links to the Nutrient Definition File by Nutr_No
- Links to the Source Codes File by Src_Cd
- Links to the Gram Weight File by NDB No.

Table 2 - Nutrient Data File: (File Name = NUT_DATA)

Field Name	Type	Blank	Description
NDB_No	A 5*	N	5-digit Nutrient Data Bank number.
Nutr_No	A 3*	N	3-digit unique identifier code for a nutrient
Nutr_Val	N 10.3	N	Amount in 100 grams, edible portion. (The number of decimal places displayed does not always reflect the accuracy of the data)
Sample_Ct	N 5.0	N	Number of samples, 0 if nutrient value was not analytical
Std_Error	N 8.3	Y	Standard error of the mean. Null if could not be calculated
Src_Cd	A 2	N	Code indicating type of data
Blank_1			Field reserved for future use
Blank_2			Field reserved for future use
Blank_3			Field reserved for future use

Table 3 - Nutrients included in the Nutrient Data File

Proximates	Other Food Components	Vitamins
Protein	Amino Acids	Ascorbic acid
Fat	Caffeine	Thiamin
Carbohydrate	Theobromine	Riboflavin
Moisture	Alcohol	Niacin
Ash		Pantothenic acid
Total Dietary Fiber		Vitamin B_6
Energy		Folate
	Minerals	Vitamin B_{12}
Lipids	Calcium	Vitamin A
Cholesterol	Iron	(IU and RE)
Total Saturated Fatty Acids	Magnesium	Vitamin E (a-TE)
Total Monounsaturated Fatty Acids	Phosphorus	
Total Polyunsaturated Fatty Acids	Potassium	
Individual Fatty Acids	Sodium	
Phytosterols	Zinc	
	Copper	
	Manganese	

Table 4 - Nutrient Definition File (File Name = NUTR_DEF)

The Nutrient Definition File is the support file to the Nutrient Data File. It identifies the 3 digit nutrient number code with the unit of measure, INFOODS tagname, description, and the IFDA field number.

- Links to Nutrient Data File by Nutr_No

Table 4 - Nutrient Definition File (File Name = NUTR_DEF)

Field Name	Type	Blank	Description
Nutr_No	A 3*	N	3-digit unique identifier code for a nutrient
Units	A 6	N	Units of measure - mg, g, mcg, etc.
Tagname	A 20	N	INFOODS Tagnames. A unique abbreviation for a food component developed by INFOODS to aid in the interchange of data.
NutrDesc	A 60	N	The name of the food component
IFDA_No	A 3	N	Number assigned by IFDA to each nutrient. Used to cross-reference data in the IFDA data exchange format.
Blank_1			Field reserved for future use

Table 5 - Gram Weight File: (File Name = WEIGHT)

The Gram Weight File contains the gram weight for household measures for a food item with the measure number that links it to the description of the measure.

- Links to Food Description File by NDB_No
- Links to the Measure Description File by Msre_No
- Links to Nutrient Data File by NDB_No

Table 5 - Gram Weight File: (File Name = WEIGHT)

Field Name	Type	Blank	Description
NDB_No	A 5*	N	5-digit Nutrient Data Bank No.
Msre_No	A 5*	N	A unique code in the Measure Description File referencing the description
Gm_wt	N 9.2	N	The weight of the food item
Blank_1			Field reserved for future use

Table 6 - Measure Description File: (File Name = MEASURE)

The Measure Description File is the support file for the Gram Weight File. It contains the 5 digit measure number and measure description.

- Links to the Gram Weight File by Msre_No

Table 6 - Measure Description File: (File Name = MEASURE)

Field Name	Type	Blank	Description
Msre_No	A 5*	N	5 digit code denoting the measure
Msre_Desc	A 120	N	The description of the measure, i.e. "cup", "cup, chopped", "tomato", "tbsp", etc.
Blank_1			Field reserved for future use

The Gram Weight and Measure Description files together allow expansion of the number and types of household weights beyond the two previously published in AH-8 or earlier releases of the database as shown in Figure 2. Using these two files one can calculate the nutrient profile of a food for the specified measures and generate reports with the measure description as headings.

Figure 2 - Example showing relationships between Gram Weight and Measure Description files for NDB No. 11252, Lettuce, Iceberg, raw.

Gram Weight File

NDB No.	Measure ID	Weight
11252	10160	55
11252	61018	755
11252	61025	15
11252	61116	20

Measure Description File

Measure ID	Measure Description
10160	1 cup, shredded or chopped
61018	1 large head
61025	1 large leaf
61116	1 leaf

Table 7 - Source Code File: (File Name = SOURCE)

- Links to the Nutrient Data File by Src_Cd

Table 7 - Source Code File: (File Name = SOURCE)

Field Name	Type	Blank	Description
Src_Cd	A 2*	N	2 digit code
SrcCd_Desc	A 60	N	Description of source code that identifies the type of nutrient data.

The Source Code File contains codes to give the user an indication of the type of data in the Nutrient Data File. In previous versions of the SR there was not a separate field to indicate the type of data in the file. The standard error field was used for this purpose. If there was a positive value in the standard error field, the nutrient value was based on analytical data. If the value with no standard error was published in the printed sections of Handbook 8, a -1 was placed in this field. If a value was missing from the printed Handbook section, but was imputed for SR, then a -4 was placed in the standard error field. For breakfast cereals, where values for added nutrients were based on the label declaration from the manufacturer, a code of -5 was placed in the field.

In converting to the new format, a value of -1 in the standard error field was converted to a 1 and moved to the new source code field. The standard error field was then blank if an actual value was not reported. The -4 was converted to 4 and the -5 was converted to 5. To improve the usability of the database, food specialists in NDL have filled in nutrient values for many proximate components, total dietary fiber, vitamin and mineral values. Values for other nutrients, such as alcohol and vitamin E, were filled in because the food items are part of the database that is used for the USDA Continuing Survey of Food Intakes by Individuals (CSFII) (28).

We have added additional source codes to be more specific about the type of data used for processed and brand name products in SR11. At this time, previous versions of SR were not reviewed to revise source codes. Therefore the new source codes that have been added are used only for items that are new or were revised for SR11. As existing items are revised, source codes will be updated.

The few exceptions are:

- Carbohydrate values of zero in all animal products were given the source code of 7 which indicates an assumed zero.
- Carbohydrate values which are calculated by difference were given a source code of 4.
- Energy values which were calculated by Atwater factors were given a source code of 4.
- Cholesterol and vitamin B_{12} values of zero in plant products were given a source code of 7.
- Vitamin C and total dietary fiber values of zero in animal products were given a source code of 7.

Table 8 - Source Code List

Code	Description
1	The value is analytical or derived from analytical.
4	The value is imputed.
5	The value upon which a manufacturer based their label claim for added nutrients (Used primarily for Breakfast Cereals and Infant Formulas)
7	The value is an assumed zero. The nutrient is not expected to be present because biologically it could not be present, such as dietary fiber in animal products, or the nutrient is expected to be present in only insignificant amounts, such as vitamin C in meat products.
8	The value is calculated from the nutrient label by NDL.
9	The value is calculated by the manufacturer, not adjusted or rounded for NLEA compliance.
12	The value is analytical, supplied by the manufacturer with partial documentation.

Table 9 - Food Group Description File: (File Name = FD_GROUP)

- Links to the Food Description File by FdGp_Cd

Table 9 - Food Group Description File: (File Name = FD_GROUP)

Field Name	Type	Blank	Description
FdGp_Cd	A 4*	N	Four digit code identifying a food group. Currently only the first 2 digits are assigned. In the future the last 2 digits may be utilized.
FdGp_Desc	A 60	N	The name of the food group

Abbreviated File

This file is an adaptation of the Abbreviated File included with earlier releases and is provided as a convenience for users of that file. It does not contain the full complement of nutrients as the relational files described above. It contains NDB No., the short description, 32 nutrients, and two household measures. Because of the restructuring of the SR files, some changes were made to this file as well: 1) The 20-character name is replaced with the 60-character short description; 2) the nutrients magnesium, zinc, copper, manganese, vitamin B_6, pantothenic acid, folate, vitamin B_{12} , and vitamin E as mg a-tocopherol equivalents have been added; and 3) only the first two weights and their description for each NDB No. in the gram weight file are included, which may not be the same two weights as in previous releases of this file. To obtain additional information this file can be linked to the other files listed above by the NDB No.

Table 10 - Layout of Abbreviated File

Field Name	Type	Description
NDB No.	A 5*	5-digit Nutrient Data Bank number.
Shrt_Desc	A 60	60 Character abbreviated description of food item. The 200 character description and other descriptive information can be obtained by linking to the Food Description File.
Water	N 10.3	Water in grams per 100 g
Energ_Kcal	N 10.3	Food Energy in kilocalories per 100 g
Protein	N 10.3	Protein in grams per 100 g
Tot_Lipid	N 10.3	Total lipid (fat) in grams per 100 g
Carbohydrt	N 10.3	Carbohydrate, by difference in grams per 100 g
Fiber_TD	N 10.3	Total dietary fiber in grams per 100 g
Ash	N 10.3	Ash in grams per 100 g
Calcium	N 10.3	Calcium in milligrams per 100 g
Phosphorus	N 10.3	Phosphorus in milligrams per 100 g
Iron	N 10.3	Iron in milligrams per 100 g
Sodium	N 10.3	Sodium in milligrams per 100 g
Potassium	N 10.3	Potassium in milligrams per 100 g
Magnesium	N 10.3	Magnesium in milligrams per 100 g
Zinc	N 10.3	Zinc in milligrams per 100 g
Copper	N 10.3	Copper in milligrams per 100 g
Manganese	N 10.3	Manganese in milligrams per 100 g
Vit_A	N 10.3	Vitamin A in IU per 100 g
Vit_E	N 10.3	Vitamin E in mg a-tocopherol equivalents
Thiamin	N 10.3	Thiamin in milligrams per 100 g
Riboflavin	N 10.3	Riboflavin in milligrams per 100 g
Niacin	N 10.3	Niacin in milligrams per 100 g
Panto_acid	N 10.3	Pantothenic acid in milligrams per 100 g
Vit_B6	N 10.3	Vitamin B_6 in milligrams per 100 g
Folate	N 10.3	Folate in micrograms per 100 g

Table 10 - Layout of Abbreviated File (continued)

Field Name	Type	Description
Vit_B12	N 10.3	Vitamin B_{12} in micrograms per 100 g
Vit_C	N 10.3	Vitamin C in milligrams per 100 g
FA_Sat	N 10.3	Saturated fatty acid in grams per 100 g
FA_Mono	N 10.3	Monounsaturated fatty acids in grams per 100 g
FA_Poly	N 10.3	Polyunsaturated fatty acids in grams per 100 g
Cholestrl	N 10.3	Cholesterol in milligrams per 100 g
GmWt_1	N 9.2	The first household weight for this item from the Gram Weight File. For the complete list and description of the measure, link to that file.
GmWt_Desc1	A 120	Description of household weight number 1
GmWt_2	N 9.2	The second household weight for this item from the Gram Weight File. For the complete list and description of the measure, link to that file.
GmWt_Desc2	A 120	Description of household weight number 2
Refuse_Pct	N 2.0	The percent refuse. For description of refuse, link to the Food Description File

References

1. International Food Distributors Association. 1995. IFDA Product Data Exchange Format, Version 3.0. Washington, D.C.
2. Klensin, J.C., D. Feskanich, V. Lin, A.S. Truswell, and D.A.T. Southgate. 1989. Identification of Food Components for INFOODS Data Interchange. United Nations University. Tokyo. 106 pp.
3. U.S. Department of Agriculture, Agricultural Research Service. 1976. Composition of Foods: Dairy and Egg Products; Raw, Processed, Prepared. Agric. Handb. No. 8-1, 144 pp.
4. U.S. Department of Agriculture, Agricultural Research Service. 1977. Composition of Foods: Spices and Herbs; Raw, Processed, Prepared. Agric. Handb No. 8-2, 51 pp.
5. U.S. Department of Agriculture, Agricultural Research Service. 1978. Composition of Food: Baby Foods; Raw, Processed, Prepared. Agric. Handb. No. 8-3, 231 pp
6. U.S. Department of Agriculture, Agricultural Research Service. 1979. Composition of Foods: Fats and Oils; Raw, Processed, Prepared. Agric. Handb. No. 8-4, 142 pp.
7. U.S. Department of Agriculture, Agricultural Research Service. 1979. Composition of Foods: Poultry Products; Raw, Processed, Prepared. Agric. Handb. No. 8-5, 330 pp.
8. U.S. Department of Agriculture, Agricultural Research Service. 1980. Composition of Foods: Soups, Sauces, and Gravies; Raw, Processed, Prepared. Agric. Handb. No. 8-6, 228 pp.
9. U.S. Department of Agriculture, Agricultural Research Service. 1980. Composition of Foods: Sausages and Luncheon Meats; Raw, Processed, Prepared. Agric. Handb. No. 8-7, 92 pp.
10. U.S. Department of Agriculture, Agricultural Research Service. 1982. Composition of

Foods: Breakfast Cereals; Raw, Processed, Prepared. Agric. Handb. No. 8-8, 160 pp.
11. U.S. Department of Agriculture, Agricultural Research Service. 1982. Composition of Foods: Fruits and Fruit Juices; Raw, Processed, Prepared. Agric. Handb. No. 8- 9, 283 pp.
12. U.S. Department of Agriculture, Agricultural Research Service. 1992. Composition of Foods: Pork Products; Raw, Processed, Prepared. Agric. Handb. No. 8-10, 223 pp.
13. U.S. Department of Agriculture, Agricultural Research Service. 1984. Composition of Foods: Vegetables and Vegetable Products; Raw, Processed, Prepared. Agric. Handb. No. 8-11, 502 pp.
14. U.S. Department of Agriculture, Agricultural Research Service. 1984. Composition of Foods: Nut and Seed Products; Raw, Processed, Prepared. Agric. Handb. No. 8-12, 137 pp.
15. U.S. Department of Agriculture, Agricultural Research Service, 1990. Composition of Foods: Beef Products; Raw, Processed, Prepared. Agric. Handb. No. 8-13, 412 pp.
16. U.S. Department of Agriculture, Agricultural Research Service, 1986. Composition of Foods: Beverages; Raw, Processed, Prepared. Agric. Handb. No. 8-14, 173 pp.
17. U.S. Department of Agriculture, Agricultural Research Service, 1987. Composition of Foods: Finfish and Shellfish Products; Raw, Processed, Prepared. Agric. Handb. No. 8-15, 192 pp.
18. U.S. Department of Agriculture, Agricultural Research Service, 1986. Composition of Foods: Legumes and Legume Products; Raw, Processed, Prepared. Agric. Handb. No. 8-16, 156 pp.
19. U.S. Department of Agriculture, Agricultural Research Service, 1989. Composition of Foods: Lamb, Veal, and Game Products; Raw, Processed, Prepared. Agric. Handb. No. 8-17, 251 pp.
20. U.S. Department of Agriculture, Agricultural Research Service, 1992. Composition of Foods: Baked Products; Raw, Processed, Prepared. Agric. Handb. No. 8-18, 467 pp.
21. U.S. Department of Agriculture, Agricultural Research Service, 1991. Composition of Foods: Snacks and Sweets; Raw, Processed, Prepared. Agric Handb. No. 8-19, 341 pp.
22. U.S. Department of Agriculture, Agricultural Research Service, 1989. Composition of Foods: Cereal Grains and Pasta; Raw, Processed, Prepared. Agric. Handb. No. 8-20, 137 pp.
23. U.S. Department of Agriculture, Agricultural Research Service, 1988. Composition of Foods: Fast Foods; Raw, Processed, Prepared. Agric. Handb. No. 8-21, 194 pp.

U.S. Department of Agriculture, Agricultural Research Service, Composition of Foods. Raw, Processed, Prepared Agric. Handb. No. 8.
24. 1989 Supplement, 1990. 96 pp.
25. 1990 Supplement, 1991. 114 pp.
26. 1991 Supplement, 1992. 85 pp.
27. 1992 Supplement, 1993. 144 pp.

28. U.S. Department of Agriculture. 1995. CSFII/DHKS Data Set, Documentation and Technical Support Files: The 1994 Continuing Survey of Food Intakes by Individuals and the 1994 Diet Health and Knowledge Survey. National Tech. Info. Service. Springfield, VA. CD-ROM: Accession No. PB95-501010.